PATENT
Docket No. 235.00040201

RUBREDOXIN FUSION PROTEINS, PROTEIN EXPRESSION SYSTEM AND METHODS

This application claims the benefit of U.S. Provisional Application Serial No. 60/114,034, filed December 29, 1998.

Field of the Invention

The invention relates to a fusion protein comprising a fusion partner, in this case rubredoxin, fused directly or indirectly to a protein or peptide of interest, together with methods and materials for producing the fusion protein in a host cell and purifying the fusion protein. The fusion protein can, in some embodiments of the invention, be cleaved to release the peptide or protein of interest for further use or analysis. The invention further relates to immunogenic compounds comprising a rubredoxin as a carrier molecule linked to an antigen or a hapten.

Background of the Invention

The recombinant production of biologically active peptides and proteins in *E. coli* currently offers an attractive alternative to chemical synthesis. This is especially true in the case of longer chain peptides (e.g., longer than about 30-35 amino acids), very hydrophobic peptides, and peptides containing cysteines which depend on proper folding for solubility and activity. However, synthesis of peptides in *E. coli* is not without problems. Foreign peptides may, for example, be susceptible to proteolytic degradation. Additionally, incorrect folding of proteins and/or aggregation of hydrophobic proteins into inclusion bodies can cause insolubility, necessitating the use of chaotropic agents like 8M urea, 6M guanidine hydrochloride and, in extreme cases, guanidine thiocyanate

to recover them. It can be very difficult to restore biological activity to the protein or peptide after treatment with these solubilizing agents.

Some of the strategies employed to overcome the problems of protein stability and solubility in *E. coli* include the use of fusion partners such as maltose binding protein (31 kD) (P. Riggs, in Ausebel, F.M. et al. (Eds) *Current Protocols in Molecular Biology*, Greene Associates/Wiley Interscience, N.Y. (1990)), thioredoxin (U.S. Pat. No. 5,646,016, issued Jul. 8, 1997; U.S. Pat. No. 5,270,181, issued Dec. 14, 1993; U.S. Pat. No. 5,292,646, issued Mar. 8, 1994) and glutathione-S-transferase (28kD) (D. Smith et al., *Gene* 67: 31-40 (1988); U.S. Pat. No. 5,654,176); and the use of protease deficient strains of *E. coli* (Bibi et al., *Proc. Nat'l. Acad. Sci. (USA)* 90 :9209 (1993); D. Alexander et al., *Protein Exp. Purif.*, 3:204 (1992)). The importance of the cellular redox environment as a factor affecting folding and solubility of foreign proteins has been demonstrated through the use of the redox-active protein thioredoxin (12kD) as a fusion partner in expression systems (E. Lavallie et al., *Biotechnology* 11:18 (1993)) and through the synthesis of proteins in thioredoxin reductase (trx-) negative strains of *E. coli* (A. Darman et al., *Science* 262:1744 (1993)). These fusion systems have proven very useful, but the fusion products are sometimes difficult to follow during purification and there is still no assurance that any given protein will fold properly and/or become or remain soluble in any of the fusion systems in current use. Moreover, although the fusion partners maltose binding protein, glutathione-S-transferase and thioredoxin are typically derived from bacteria or protozoa, the existence of closely related mammalian and avian analogues of these fusion partners makes them unsuitable for use as anchor proteins for haptens in antibody production or in vaccines. Thus, continued development of new protein expression systems based on recombinant protein fusions with a stable carrier is necessary to advance the art of recombinant protein production.

Summary of the Invention

The present invention provides a recombinant rubredoxin fusion protein containing an N-terminal rubredoxin constituent and a C-terminal fused polypeptide. The fusion protein is capable of binding Fe^{2+} when properly folded, giving it a red color that makes it easy to follow during purification. The N-terminal rubredoxin constituent of the rubredoxin fusion protein preferably contains a rubredoxin obtained from an anaerobic bacterium, more preferably *Desulfovibrio vulgaris*, or a biologically active analogue, fragment, or modification thereof. Advantageously, the C-terminal fused polypeptide can be a polypeptide that is insoluble or known to form inclusion bodies in a host cell. For example, amyloid peptide, leptin, proinsulin, trypsin inhibitor, and the extracellular domain of luteinizing hormone receptor, including biologically active fragments, modifications and analogues thereof, can be fused to rubredoxin to yield rubredoxin fusion proteins of the invention. The linkage between the N-terminal rubredoxin constituent and C-terminal fused polypeptide can, but need not, be a cleavable linkage.

Antigenic or immunogenic rubredoxin fusion proteins of the invention have C-terminal fused polypeptides that are antigens (including polyfusion antigens) or haptens. The rubredoxin constituent serves as the carrier molecule to yield an immunogenic fusion product. Because rubredoxin itself is only negligibly antigenic, there is no need to include in the antigenic or immunogenic fusion protein a cleavage site to allow cleavage of the N-terminal rubredoxin constituent from C-terminal fused polypeptide. The invention includes a method for producing an antibody to a C-terminal fused polypeptide by eliciting in a host cell, preferably a mammalian host cell, an immune response to a rubredoxin fusion protein containing the C-terminal fused polypeptide. The antibodies thus generated can be polyclonal or monoclonal, and are preferably not, but can be, cross-reactive with rubredoxin. The invention further provides a polypeptide vaccine containing an antigenic or immunogenic rubredoxin fusion protein of the invention, and a polynucleotide vaccine containing a polynucleotide encoding an antigenic or immunogenic rubredoxin fusion protein.

The N-terminal rubredoxin constituent of the rubredoxin fusion protein can be directly or indirectly linked to the C-terminal fused polypeptide. In embodiments in which the linkage is indirect, the fusion protein contains a spacer region positioned between the N-terminal rubredoxin constituent and the C-terminal fused polypeptide. This intervening spacer region optionally contains a proteolytic cleavage site, an affinity purification sequence, or both. Alternatively, the N-terminal rubredoxin constituent can be directly linked to the C-terminal fused polypeptide, with no intervening spacer region.

The present invention further provides a recombinant polynucleotide having a nucleotide sequence that encodes a rubredoxin fusion protein as described herein. In addition, the invention includes an expression vector that contains a promoter operably linked to a nucleotide sequence encoding a rubredoxin fusion protein, and a host cell transformed with an expression vector comprising a recombinant polynucleotide comprising a nucleotide sequence encoding a rubredoxin fusion protein. Preferably the host cell is a bacterial cell.

Also provided by the invention is an expression vector that contains a nucleotide sequence encoding rubredoxin or a biologically active analogue, fragment, or modification thereof; an intervening nucleotide sequence encoding a spacer region; and a multiple cloning region that contains at least one restriction endonuclease recognition site. The intervening nucleotide sequence preferably includes all or a portion of the multiple cloning region, and the spacer region encoded by the intervening nucleotide sequence preferably contains at least one of one of a proteolytic cleavage site and an affinity purification sequence. A preferred expression vector is pRUBEX3.

The invention further provides a method for making a rubredoxin fusion protein that involves introducing into a host cell a recombinant polynucleotide having a nucleotide sequence encoding a rubredoxin fusion protein, followed by expressing the fusion protein in the host cell. Optionally, the fusion protein is removed from the host cell and further purified as desired. Optionally, the fusion protein contains an affinity purification sequence that

permits reversible binding of the fusion protein to an affinity chromatography matrix thereby facilitating removal of contaminants.

The invention also provides a recombinant method for making a polypeptide that includes introducing into a host cell a recombinant polynucleotide having a nucleotide sequence encoding a rubredoxin fusion protein; expressing the fusion protein in the host cell; removing the fusion protein from the host cell; and cleaving the fusion protein to yield the rubredoxin constituent and the polypeptide. Optionally, this method further includes separating the polypeptide from the rubredoxin constituent after cleavage.

Brief Description of the Drawings

Figure 1 depicts (a) a schematic of the vector pRUBEX3, including the Multiple Cloning Region (MCR); and (b) the nucleotide sequence (SEQ ID NO:1) of a portion of pRUBEX3 together with the amino acid sequence encoded thereby (SEQ ID NO:2) wherein the 52 amino acids of rubredoxin (SEQ ID NO:3) are underlined; the amino acids of the polyhistidine (polyHis) sequence (i.e., His-His-His-His-His-His) (SEQ ID NO:4) are in bold; the eight amino acids of the flag peptide are double-underlined (DYKDDDDK; i.e., Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys) (SEQ ID NO:5); the five amino acids of the enterokinase site (DDDDK; i.e., Asp-Asp-Asp-Asp-Lys) (SEQ ID NO:6) are in bold and double-underlined; and the restriction sites are labeled and in italics. Another embodiment of pRUBEX3 (not pictured) includes, in place of the polyhistidine sequence, the affinity tag His-Gly-Leu-His (SEQ ID NO:7).

Figure 2 shows a portion of the nucleotide sequence (SEQ ID NO:8) and the encoded amino acid sequence (SEQ ID NO:9) for the $A\beta_{1-42}$ rubredoxin fusion construct; the underlined amino acid sequence (SEQ ID NO: 10) represents the $A\beta_{1-42}$ peptide and the intervening spacer region comprises a flag peptide sequence (SEQ ID NO:5), a polyhistidine (polyHis) sequence for use in affinity purification (SEQ ID NO:4), and the amino acid sequence IEGR (in bold) (i.e., Ile-Glu-Gly-Arg) (SEQ ID NO:11), which is the recognition site for the restriction protease Factor Xa. Another embodiment of the $A\beta_{1-42}$ rubredoxin

fusion construct (not pictured) includes, in place of the polyhistidine sequence, the affinity tag His-Gly-Leu-His (SEQ ID NO:7).

Figure 3 is a schematic of the expression vector pRUBEX2-LHR, which contains the amino-terminal 298 amino acid residues of human luteinizing hormone receptor (LHR), representing the extracellular domain, cloned into the *Nde*I/*Bam*HI site of pRUBEX2; the resulting construct encodes a fusion protein consisting of rubredoxin followed by a spacer region comprising a polyhistidine tag to facilitate purification of the fusion protein and a Factor Xa recognition site that directly precedes the LHR coding region. Another embodiment of pRUBEX2-LHR (not pictured) includes, in place of the polyhistidine sequence, the affinity tag His-Gly-Leu-His (SEQ ID NO:7).

Figure 4 is a schematic of the expression vector pRUBEX1-LHR, which contains cDNA encoding the amino-terminal 340 amino acids of human luteinizing hormone receptor (LHR), representing the extracellular domain, cloned into the *Bam*HI site of pRUBEX1; the resulting construct encodes a fusion protein consisting of the N-terminal extracellular domain of human LHR directly fused to the carrier protein rubredoxin at the C-terminal end of rubredoxin.

Figure 5 shows Tris-tricine gel electrophoresis of rubredoxin fusion proteins and digestion products.

Figure 6 is a Western-blot analysis of purified pig leptin/rubredoxin fusion protein and a Factor Xa digest of the fusion protein.

Detailed Description

Rubredoxin is an electron carrier protein originally isolated and then cloned from the anaerobic sulfate reducing bacteria, *Desulfovibrio vulgaris*. Since then, rubredoxins from several different anaerobic organisms have been discovered and characterized. Rubredoxin is a small redox protein (5.6 kD) carrying a single non-haem iron center. The crystal structure of the protein has been solved and reveals a free carboxy-terminal end, making it well-suited for fusing peptides. The iron center imparts a red color to the protein (absorption

maxima at 390 nm and 495 nm) providing a visible marker for easy monitoring during purification protocols. The red color also serves indicator as to whether the fusion protein has folded correctly, since an incorrectly folded protein will not bind the metal. Recombinant rubredoxin can be produced at high levels (50-60 mg/L of purified protein) in *E. coli* and is very soluble, biologically active and stable. Conveniently, rubredoxin is a thermostable protein and can withstand 70°C-80°C for more than an hour without denaturation. It also retains its metal center in denaturing agents like 0.5% SDS and 6M urea.

The present invention utilizes rubredoxin as a protein fusion partner in the creation of a simple, reliable, reproducible, scalable and economical recombinant protein expression system. The presence of a correctly folded fusion protein can be visually tracked during purification due to the effects of the iron atom. Moreover, proper folding of the fused protein may be facilitated by the redox functionality of rubredoxin. That is, the presence of high levels of an active, foreign electron carrier protein like rubredoxin is likely to beneficially alter the redox microenvironment of the fusion protein. Folding of a protein fused to an electron carrier protein such as rubredoxin is thus likely to be affected by the redox state of the carrier as well as the oxidation state within the cell. The protein expression system of the invention is particularly useful for producing proteins and peptides, such as β-amyloid peptide, leptin and pro-insulin, that are otherwise insoluble or tend to form inclusion bodies in recombinant systems. For example, leptin from both rat and pig are known to form inclusion bodies and require the use of chaotropic agents for solubilization, and pig leptin can be efficiently produced using the protein expression system of the present invention.

Accordingly, the invention provides a rubredoxin fusion protein and, further, a recombinant polynucleotide containing a nucleotide sequence that encodes the rubredoxin fusion protein of the invention, as well as a polynucleotide having a nucleotide sequence complementary thereto. A rubredoxin fusion protein is a protein that comprises a rubredoxin constituent and a polypeptide of interest. The rubredoxin constituent comprises the N-terminus of the fusion protein, and the fused polypeptide constitutes the C-

terminus of the fusion protein. In a preferred embodiment, the rubredoxin fusion protein contains an intervening spacer region between the rubredoxin constituent and the fused polypeptide, as described more fully below.

The rubredoxin constituent of the rubredoxin fusion protein is composed primarily of a rubredoxin polypeptide and serves as a "carrier" or "ballast" for the fused polypeptide. For example, the rubredoxin constituent can assist in stabilization, folding, solublization and/or targeting of the fused polypeptide, while providing additional options for detecting, isolating and purifying the polypeptide. In addition to a rubredoxin polypeptide (its main and often sole component), the rubredoxin constituent of the rubredoxin fusion protein optionally contains one or more of an affinity purification sequence (described below), a signal sequence or a targeting sequence, for example a sequence targeting the fusion protein to a bacterial periplasm or causing the fusion protein to be secreted into the surrounding media, which is particularly useful in eukaryotic expression systems. A signal sequence or targeting sequence is preferably located at the N-terminus of the rubredoxin fusion protein (and hence is located at the N-terminal end of the rubredoxin constituent), whereas an affinity purification sequence can be positioned at the N-terminus of the fusion protein, within the rubredoxin polypeptide sequence itself, or C-terminal to the rubredoxin polypeptide. In the latter case, the affinity purification sequence may be thought of as part of the intervening spacer region rather than part of the rubredoxin constituent per se. Inclusion of the optional affinity sequence, signal sequence and/or targeting sequence must not prevent the rubredoxin polypeptide from folding properly. Whether or not the rubredoxin polypeptide folds properly (i.e., whether or not it is biologically active) can be easily assayed by determining whether it can bind a divalent cation, particularly Fe^{2+}, as discussed in more detail below. For example, engineering a histidine tag (His-His-His-His-His-His, SEQ ID NO:4) as an affinity purification sequence at the N-terminus of the fusion protein caused the rubredoxin polypeptide to fail to bind iron. However, use of an N-terminal affinity sequence that is less highly charged could result in a rubredoxin polypeptide that does bind iron, i.e., a rubredoxin fusion protein of the invention.

The rubredoxin fusion protein is a single polypeptide chain wherein the rubredoxin constituent is linked by way of a peptide bond, either directly or indirectly, to the polypeptide of interest. This linkage is termed "direct" in embodiments of the rubredoxin fusion protein containing no intervening spacer region; it is termed "indirect" in embodiments of the rubredoxin fusion protein that contain an intervening spacer region. The fused polypeptide can have a preselected or predetermined amino acid sequence, a random amino acid sequence, or an unknown amino acid sequence. It is to be understood that the terms peptide, polypeptide, and protein as used herein are interchangeable, as the invention is not limited by the length or the function of the amino acid sequence linked to the rubredoxin constituent. As used herein these terms all refer generally to a plurality of amino acids joined together in a linear chain via peptide bonds. In some contexts, the term "peptide" may be used to connote a shorter polypeptide such as dipeptide, tripeptide, or oligopeptide; the term oligopeptide typically connoting a polypeptide having between 2 and about 50 or more amino acids. However, the term "peptide" is not limited to polypeptides of any particular length. The term "protein" is sometimes used herein to mean a functionally folded polypeptide of any length having structural, enzymatic or other active properties. Regardless of the nomenclature used, however, no limitations on the length or the function of the fused polypeptide or protein are intended.

The rubredoxin constituent of the fusion protein comprises a rubredoxin polypeptide. Preferably, the rubredoxin polypeptide has the wild-type amino acid sequence of a rubredoxin protein obtained from an anaerobic bacterium, preferably from *Desulfovibrio, Clostridium, Desulfoarculus* or *Pyrococcus* spp., more preferably from *D. vulgaris, D. vulgaris* (Hildenborough), *C. pasteurianum, C. butyricum, D. baarsii* or *P. furiosa*. GenBank Accession numbers for nucleotide sequences encoding rubredoxins include D76419 (*rub* gene for *D. vulgaris*), M28848 (*rub* gene for *D. vulgaris* (Hildenborough), M60116 (*C. pasteurianum* rubredoxin gene), Y11875 (*C. butyricum* rubredoxin gene), and X99543 for *D. baarsii*. A particularly preferred amino acid sequence for the rubredoxin polypeptide is an amino acid sequence of a rubredoxin from

D. vulgaris, more preferably SEQ ID NO:3 (Fig. 1). The amino acid sequence of the rubredoxin polypeptide useful in the fusion protein of the invention is not intended to be limited to the exact wild-type amino acid sequence of naturally occurring rubredoxin proteins; rather, the rubredoxin polypeptide includes biologically active analogues, fragments, or modifications of any and all naturally occurring rubredoxin proteins.

When used herein to describe a rubredoxin analogue, fragment, or modification thereof, the term "biologically active" means that the rubredoxin analogue, fragment or modification thereof can, when present as a component of the fusion protein of the invention, can bind a divalent cation. Preferably, biologically active rubredoxin or analogue, fragment, or modification thereof binds Zn^{2+} or Fe^{2+}; more preferably it binds Fe^{2+}. Biological activity (e.g., iron-binding activity) of a rubredoxin polypeptide can be easily assayed by simply observing the characteristic visible spectrum of a rubredoxin that has bound iron. Moreover, iron binding can be visually detected because the bound complex is red. Binding of Fe^{2+} by the fusion protein is indicative of proper folding of its rubredoxin polypeptide.

Naturally occurring rubredoxin is a small protein; for example, rubredoxin from *D. vulgaris* contains about 52 amino acids. A "fragment" of rubredoxin means a rubredoxin that has been truncated at the C-terminus; preferably, the fragment is at least about 40 amino acids in length, more preferably it is at least about 45 amino acids in length.

An "analogue" of rubredoxin means a rubredoxin that contains one or more amino acid substitutions, deletions, additions, or rearrangements. For example, it is well-known in the art of protein biochemistry that an amino acid belonging to a grouping of amino acids having a particular size or characteristic (such as charge, hydrophobicity and hydrophilicity) can often be substituted for another amino acid without altering the activity of a protein, particularly in regions of the protein that are not directly associated with biological activity. Thus, a rubredoxin polypeptide useful in a fusion protein according to the invention includes a rubredoxin that contains amino acid substitutions at sites such that the iron-binding activity of the polypeptide is not

eliminated. Substitutes for an amino acid may be selected from other members of the class to which the amino acid belongs. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and tyrosine. Polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid. Examples of preferred conservative substitutions include Lys for Arg and *vice versa* to maintain a positive charge; Glu for Asp and *vice versa* to maintain a negative charge; Ser for Thr so that a free -OH is maintained; and Gln for Asn to maintain a free NH_2. Likewise, rubredoxin polypeptides containing deletions or additions of one or more contiguous or noncontiguous amino acids that do not eliminate the biological activity of rubredoxin (i.e., iron binding) are also contemplated.

Preferably, a rubredoxin analogue has at least about 80% amino acid identity with a reference rubredoxin protein; more preferably it has at least about 90% amino acid identity with a reference rubredoxin protein. The reference rubredoxin protein is preferably a rubredoxin from *D. vulgaris*; more preferably it is SEQ ID NO:3. Amino acid identity is defined in the context of a homology comparison between the rubredoxin analogue and the reference rubredoxin protein. The two amino acid sequences are aligned in a way that maximizes the number of amino acids that they have in common along the lengths of their sequences; gaps in either or both sequences are permitted in making the alignment in order to maximize the number of shared amino acids, although the amino acids in each sequence must nonetheless remain in their proper order. The percentage amino acid identity is the higher of the following two numbers: (a) the number of amino acids that the two polypeptides have in common within the alignment, divided by the number of amino acids in the rubredoxin analogue, multiplied by 100; or (b) the number of amino acids that the two polypeptides have in common within the alignment, divided by the number of amino acids in the reference rubredoxin protein, e.g., SEQ ID NO:3, multiplied by 100.

"Modified" rubredoxin includes rubredoxins chemically or enzymatically derivatized at one or more constituent amino acid, including side chain modifications, backbone modifications, and N- and C- terminal modifications including acetylation, hydroxylation, methylation, amidation, and the attachment of carbohydrate or lipid moieties, cofactors, and the like.

Advantageously, the fused polypeptide of the rubredoxin fusion protein can be a polypeptide that has, in the past, been difficult to isolate in biologically active form using other recombinant expression systems. Such polypeptides include, for example, hydrophobic peptides, (that is, peptides that are insoluble in aqueous solutions), peptides and proteins that produce insoluble sedimentation aggregates known as "inclusion bodies" when overexpressed (e.g., amyloid peptides, such as β-amyloid 1-42 peptide and β-amyloid 1-40 peptide, leptins, including pig leptin and rat leptin, preproinsulin, trypsin inhibitor, and the extracellular domain of luteinizing hormone receptor), and those that become insoluble when present the high concentrations found in typical protein overproduction systems. The rubredoxin fusion protein of the invention, in contrast, is preferably soluble in aqueous solutions. More preferably, the rubredoxin fusion protein does not form insoluble sedimentation aggregates during recombinant overproduction of the fusion protein; that is, it remains soluble when overexpressed in the host cell. "Overexpression" in this context means expression of the rubredoxin fusion protein at a level of at least about 10 mg fusion protein per 100 mL cell extract (i.e., about 100 mg/L). If aggregates of the rubredoxin fusion protein do form, they are preferably capable of being resolubilized using a nonionic detergent to yield a fusion protein having a biologically active (i.e., iron-binding) rubredoxin constituent. Typically, it is not necessary to treat the protein aggregates with chaotropic agents such as urea or guanidium chloride, or even ionic detergents, to reconstitute a biologically active fusion protein.

The rubredoxin fusion protein of the invention, when it binds Fe^{2+}, is detectably labeled as a result of its red color. Optionally, the rubredoxin fusion protein is further detectably labeled. Preferably the detectable label is a radioisotope, a heavy isotope, or a fluorescent label. Isotope labels can be

conveniently incorporated into the fusion protein using isotopically labeled amino acids or precursor compounds during synthesis in the host cell using methods well known in the art. Examples of useful radiolabels include ^{3}H, ^{14}C and ^{35}S; useful heavy isotope labels are exemplified by ^{13}C and ^{15}N. A preferred fluorescent label is isofluorothiocyanate (IFTC), which can be chemically attached to the fusion protein following biosynthesis.

A particularly preferred embodiment of the fusion protein of the invention comprises a rubredoxin constituent fused, directly or indirectly, to an amyloid peptide. Preferably, the amyloid peptide is β-amyloid 1-40 or β-amyloid 1-42, or a biologically active analogue, modification or derivative thereof. Amyloid peptides that are isotopically labeled, as described above, are also especially useful. A biologically active β-amyloid peptide is one that retains the ability to aggregate into fibrils such as are observed in Alzheimer's plaques. For example, tyrosine at the 10 position in β-amyloid (Tyr10) can be changed to tryptophan to yield a bioactive β-amyloid peptide analogue, and the tryptophan can be detectably labeled using IFTC to generate modified bioactive peptide having a chartreuse color. Notwithstanding the above, the production of biologically inactive amyloid fusion proteins, for instance those having one or two amino acid deletions, additions or changes that reduce or eliminate aggregation activity, is useful for comparative or mechanistic studies and is also encompassed by the present invention. For example, arginine at the 5 position in β-amyloid (Arg5) can be changed to cysteine to yield a β-amyloid peptide analogue, and the cysteine can be labeled with IFTC to generate modified amyloid peptide that is less biologically active than the naturally occurring peptide.

Another preferred embodiment of the fusion protein of the invention is a fusion protein comprising a rubredoxin constituent linked, directly or indirectly, to the extracellular domain of luteinizing hormone receptor (LHR) or biologically active fragment, modification or analogue thereof.

Another embodiment of the invention that is particularly well suited for use in generating mammalian antibodies to the fused polypeptide is a rubredoxin fusion protein comprising an N-terminal rubredoxin constituent

directly linked to a C-terminal fused polypeptide antigen or hapten. A hapten is a low-molecular weight compound that reacts specifically with an antibody but does not stimulate antibody production (i.e., is not antigenic) unless complexed with a carrier protein. Linking the carrier protein (i.e., rubredoxin) to the hapten produces an immunogen that stimulates antibody production against the hapten. The hapten portion of the immunogenic rubredoxin fusion protein is preferably at least about four amino acids in length, more preferably at least about six amino acids in length, most preferably at least about eight amino acids in length, and is preferably less than about 50 amino acids in length, more preferably less than about 35 amino acids in length, most preferably less than about 25 amino acids in length.

One type of polypeptide antigen that is advantageously linked to the rubredoxin constituent in this embodiment of the rubredoxin fusion protein is a protein that would be insoluble or form inclusion bodies in the absence of a rubredoxin carrier. Alternatively, the polypeptide antigen portion of the rubredoxin fusion protein can contain more than one antigenic epitope fused in tandem, forming what is known as a polyfusion antigen. Rubredoxin has a significant advantage over other known carrier proteins for antibody production (such as thioredoxin, glutathione sulfotransferase and maltose binding protein) in that rubredoxins are never present in mammalian systems. Any anti-rubredoxin that is generated in the host will not cross-react with cell extracts from eukaryotic organisms. Moreover, in initial experiments in rabbits, rubredoxin has shown undetectable levels of antigenicity itself, the immune response thus being mounted against the fused peptide. However, in mammalian systems where rubredoxin may prove more antigenic, its desirability as a fusion partner could well be enhanced due to increased stimulation of the host's immune system. There is in any event no need to include in the fusion protein a cleavage site between the rubredoxin polypeptide and the fused polypeptide, since presence of the rubredoxin polypeptide does not interfere with antibody generation. In addition, there is no need to include in the fusion protein an affinity purification sequence, since the fusion product can be isolated by electrophoresis, excised from the gel, homogenized and injected directly into the host using well-known

laboratory procedures and techniques for raising mammalian or avian antibodies. The corresponding recombinant polynucleotide encoding this embodiment of the rubredoxin fusion protein includes, in the 5' to 3' direction, a nucleotide sequence encoding the rubredoxin constituent directly followed by an in-frame nucleotide sequence encoding the fused polypeptide. Notwithstanding anything above to the contrary, however, a rubredoxin fusion protein comprising a fused polypeptide antigen can, if desired, contain one or both of a cleavage site between the rubredoxin polypeptide and the fused polypeptide antigen, and an affinity purification sequence.

For other applications and uses, including, for example, large-scale protein expression, a preferred embodiment of the invention includes a rubredoxin fusion protein comprising a rubredoxin constituent that is linked indirectly to the fused polypeptide. In this embodiment of the invention, the rubredoxin fusion protein comprises an intervening spacer region positioned between the rubredoxin constituent and the fused polypeptide. The invention is not to be limited by any particular upper limit on the size of the spacer region. The optimal length of the spacer region depends on the nature of the fused peptide and can be readily determined by one of skill in the art. For example, where the spacer region contains a cleavage site, the optimal length of the spacer region can be determined by analyzing the efficiency of cleavage in test fusion proteins having spacer regions of varying lengths. Preferably, the intervening spacer region consists of less than about 100 amino acids.

In rubredoxin/β-amyloid fusion proteins made according to the invention, the spacer region preferably contains between 0 and about 100 amino acids, more preferably between about 10 and about 60 amino acids, more preferably between about 20 and about 40 amino acids. For example, in the embodiment of the invention shown in Fig. 2, the intervening space region (MHGGSEFENHHHHHHNDYKDDDDKDLIEGR (i.e., Met-His-Gly-Gly-Ser-Glu-Phe-Glu-Asn-His-His-His-His-His-His-Asn-Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys-Asp-Leu-Ile-Glu-Gly-Arg, SEQ ID NO:12) for the rubredoxin/β-amyloid fusion protein consists of 30 amino acids. An analogous intervening

spacer region that includes a His-Gly-Leu-His (SEQ ID NO:7) affinity tag contains 28 amino acids.

The intervening spacer region optionally comprises one or more proteolytic cleavage sites, one or more affinity purification sequences, and/or one or more amino acids that happen to be encoded by that portion of the multiple cloning region of the vector positioned between the nucleotide sequence encoding the rubredoxin constituent and nucleotide sequence encoding the fused polypeptide, as described in more detail below.

The proteolytic cleavage site allows enzymatic or chemical cleavage of the fusion protein into two portions, permitting separation of the fused polypeptide from the rubredoxin constituent. Thus, it must be positioned in between the rubredoxin constituent and the fused polypeptide to serve its intended function. Preferably, it is positioned at the end of the intervening spacer region so as to minimize the attachment of additional amino acids to the fused polypeptide. Chemical cleavage can be achieved, for example, by cyanogen bromide or hydroxylamine. For example, a cleavage site that comprises methionine allows cleavage to release the polypeptide of interest upon contact of the rubredoxin fusion protein with cyanogen bromide. Care must be taken with hydroxylamine as it can be relatively nonspecific under some conditions. Enzymatic cleavage can be facilitated by including as a cleavage site an amino acid sequence recognized by a restriction protease, also called an endoprotease. For example, cleavage sites recognized by thrombin, Factor Xa, renin, or enterokinase can be utilized. Preferably, cleavage of the rubredoxin fusion protein at the cleavage site yields a polypeptide having no extraneous, unintended or non-native N-terminal amino acids. To that end, the use of cleavage sites comprising Ile-Glu-Gly-Arg, SEQ ID NO:11 (IEGR, the amino acid sequence recognized by Factor Xa) or methionine (provided the second peptide component has no internal methionines), contiguous to the fused polypeptide are particularly preferred.

An affinity purification sequence is an amino acid sequence designed to facilitate purification of the fusion peptide using affinity chromatography. For example, a polyhistidine (SEQ ID NO:4) or His-Gly-Leu-

His (SEQ ID NO:7) site or "tag" can be engineered into the fusion protein to allow purification of the fusion protein using Ni-chelating affinity chromatography (commercially available from numerous sources, for example Qiagen, Boehringer Mannheim Biochemicals, and Novagen). As another example, an affinity purification system commercially available from IBI Kodak (Rochester, NY) utilizes the "flag" peptide (YKDDDDK, i.e., Tyr-Lys-Asp-Asp-Asp-Asp-Lys, SEQ ID NO:13) and a monoclonal antibody-linked resin (IGM2) that is highly specific for that peptide.

As yet another example, a chitin-binding tag can be combined with a self-cleaving protein splicing element (an intein) to permit purification of the rubredoxin fusion protein and cleavage of the fused polypeptide in a single chromatographic step. Such as system is commercially available as the IMPACT-CN system from New England BioLabs (Beverly, MA). The fusion protein binds to a chitin column. Subsequently, in the presence of a disulfide reducing agent such as dithiothreitol, β-mercaptoethanol or cysteine, the intein undergoes specific self-cleavage which releases the fused polypeptide from the chitin-bound intein tag. As discussed above, an affinity purification sequence can be positioned at essentially any location along the length of the rubredoxin fusion protein as long as it does not prevent the rubredoxin polypeptide from folding properly.

The recombinant polynucleotide of the invention includes a nucleotide sequence encoding the rubredoxin fusion protein of any of the various embodiments described above. Thus, the recombinant polynucleotide encodes, in a 5' to 3' direction, a rubredoxin constituent linked, directly or indirectly, to a polypeptide of interest; alternatively it encodes, in the 5' to 3' direction, a polypeptide of interest linked, directly or indirectly, to a rubredoxin constituent. It optionally encodes an intervening spacer region, one or more affinity sites, cleavage sites, targeting sites, and the like, as described generally for the rubredoxin fusion protein.

The invention further provides an expression vector capable of directing expression of a rubredoxin fusion protein in a host cell. The expression vector can be circular or linear, single-stranded or double stranded,

and can include DNA, RNA, or any modification or combination thereof. The vector can be a plasmid, a viral vector or a cosmid. Selection of a vector or plasmid backbone depends upon a variety of desired characteristics in the resulting construct, such as a selection marker, plasmid reproduction rate, and the like. Suitable plasmids for expression in *E. coli*, for example, include pUC(X), pKK223-3, pKK233-2, pTrc99A, and pET-(X) wherein (X) denotes a vector family in which numerous constructs are available. pUC(X) vectors can be obtained from Pharmacia Biotech (Piscataway, NH) or Sigma Chemical Co. (St. Louis, MO). pKK223-3, pKK233-2 and pTrc99A can be obtained from Pharmacia Biotech. pET-(X) vectors can be obtained from Promega (Madison, WI) Stratagene (La Jolla, CA) and Novagen (Madison, WI). To facilitate replication inside a host cell, the vector preferably includes an origin of replication (known as an "ori") or replicon. For example, ColE1 and P15A replicons are commonly used in plasmids that are to be propagated in *E. coli*.

The expression vector preferably takes the form of a DNA molecule containing a nucleotide sequence encoding the rubredoxin fusion protein of the invention, and optionally includes a promoter sequence operably linked to the coding sequence. A promoter is a DNA fragment that facilitates transcription of genetic material. Transcription is the formation of an RNA chain in accordance with the genetic information contained in the DNA. The invention is not limited by the use of any particular promoter, and a wide variety are known. Promoters act as regulatory signals that bind RNA polymerase in a cell to initiate transcription of a downstream (3' direction) coding sequence. A promoter is "operably linked" to a nucleotide sequence if it does, or can be used to, control or regulate transcription of that nucleotide sequence. The promoter used in the invention can be a constitutive or an inducible promoter. It can be, but need not be, heterologous with respect to the host cell. Preferred promoters for bacterial transformation include *lac*, *lac*UV5, *tac*, *trc*, T7, SP6 and *ara*.

The expression vector optionally includes a Shine Dalgarno site (e.g., a ribosome binding site), and a start site (e.g., the codon ATG) to initiate translation of the transcribed message to produce the enzyme. It can also include a termination sequence to end translation. A termination sequence is typically a

codon for which there exists no corresponding aminoacetyl-tRNA, thus ending polypeptide synthesis. The expression vector optionally further includes a transcription termination sequence. The rrnB terminators, which is a stretch of DNA that contains two terminators, T1 and T2, is the most commonly used terminator that is incorporated into bacterial expression systems (J. Brosius et al., *J. Mol. Biol.*, 148:107-127 (1981)).

The expression vector optionally includes one or more marker sequences, which typically encode a gene product, usually an enzyme, that inactivates or otherwise detects or is detected by a compound in the growth medium. For example, the inclusion of a marker sequence can render the transformed cell resistant to an antibiotic, or it can confer compound-specific metabolism on the transformed cell. Examples of a marker sequence are sequences that confer resistance to kanamycin, ampicillin, chloramphenicol and tetracycline.

In an alternative embodiment, the expression vector comprises a nucleotide sequence encoding a rubredoxin polypeptide and a multiple cloning region for the insertion of a polypeptide of interest. The multiple cloning region comprises at least one restriction site and preferably comprises a multiplicity of restriction sites (see, for Example, Fig. 1 showing the multiple cloning region of pRUBEX3). The multiple cloning region (sometimes referred to as a polyclonal site) is positioned such that cloning a nucleotide sequence encoding a polypeptide of interest into that site will permit expression of a rubredoxin fusion protein comprising the polypeptide of interest; for example, the polypeptide of interest will be in frame with respect to the rubredoxin constituent and the intervening spacer region, if it is present. Preferably, the expression vector comprises a nucleotide sequence encoding rubredoxin or a biologically active analogue, fragment, or modification thereof, an intervening nucleotide sequence, and a multiple cloning region comprising a multiplicity of restriction endonuclease recognition site. The intervening nucleotide sequence preferably encodes at least one of a proteolytic cleavage site and an affinity purification sequence.

Examples of expression vectors include pRUBEX1, in which the coding sequence for *D. vulgaris* rubredoxin and the fused polypeptide are directly

linked; i.e., there is no intervening spacer region between the two components; pRUBEX2, which contains an intervening spacer region comprising a histidine tag and a Factor Xa cleavage site; and pRUBEX3, which, in addition to the histidine tag and a Factor Xa cleavage site of pRUBEX2, contains as part of the intervening spacer a portion of a multiple cloning region to facilitate cloning of the nucleotide sequence encoding the fused polypeptide into the vector. Recently, pRUBEX3 has been modified to include the affinity tag His-Gly-Leu-His (SEQ ID NO:7) in place of His_6 (SEQ ID NO:4); pRUBEX3 thus modified is the most preferred expression vector.

The invention also provides a method for making a rubredoxin fusion protein. An expression vector as described above that contains a nucleotide sequence capable of directing expression of a rubredoxin fusion protein is introduced into a host cell and the rubredoxin fusion protein is then expressed in the transformed cell. Any suitable host cell can be used, without limitation. Preferably the expression vector is a DNA molecule that comprises a nucleotide sequence encoding the rubredoxin fusion protein. If the expression vector comprises RNA, as in a retroviral vector, the host cell preferably comprises a reverse transcriptase enzyme in order to facilitate expression of the rubredoxin fusion protein. Viral vectors are especially useful in eukaryotic protein expression systems, which facilitate protein glycosylation. Optionally, the fusion protein can be removed from the transformed host cell and purified. If desired, the rubredoxin fusion protein can be labeled with a radioisotope such as ^{3}H, ^{13}C, ^{15}N or ^{35}S during synthesis using methods well-known in the art.

The host cell in which the rubredoxin fusion protein is expressed in accordance with the present invention can be a bacterium, a protozoan, or a eukaryotic cell. Eukaryotic cells include, for example, plant cells and animal cells, including for example mammalian cells, yeast cells and insect cells. In methods that involve making the protein in a eukaryotic host cell, the fusion protein is preferably targeted to the endoplasmic recticulum. Suitable host cells can be differentiated or undifferentiated, and include cells growing in mammalian tissue culture, including hybridoma cells. Particularly suitable host cells are those that have been used in other protein expression systems, such as

E. coli, *Bacillus* spp., and *Streptomyces* spp. Methods of introducing expression vectors into host cells are well-known in the art; electroporation is preferred.

Rubredoxin fusion proteins that contain a polyhistidine (SEQ ID NO:4) or His-Gly-Leu-His (SEQ ID NO:7) tag can be purified by Ni-chelating chromatography. Imidazole can be used to elute the fusion protein. Typically, purification can be achieved at moderate temperatures using a single affinity chromatographic step. Ni-chelating chromatography can be performed at temperatures from about 4°C to about 60°C, depending on the thermal stability of the fused polypeptide; typically the process is performed at room temperature or colder temperatures. Optionally, the affinity chromatography can be followed with high performance liquid chromatography for further purification of the fusion proteins.

The invention further provides a method for making a polypeptide using the protein expression system described herein. A rubredoxin fusion protein comprising a cleavage site is expressed in a host cell as described herein, then removed from the host cell. Optionally, the rubredoxin fusion protein can be affinity purified at this point, if it also contains an affinity purification sequence. The polypeptide of interest is then chemically or enzymatically cleaved away from the rubredoxin constituent of the fusion protein. A preferred cleavage site comprises Ile-Glu-Gly-Arg (IEGR, SEQ ID NO:11) and the restriction protease Factor Xa is used to cleave the fusion protein to obtain the free polypeptide. The free polypeptide can be further purified away from the rubredoxin constituent by reverse phase chromatography, typically at about pH 6 to about pH 8.5, depending on the stability of the polypeptide to acid and base. In the case of β-amyloid peptides, reverse phase chromatography is preferably carried out at temperatures between about 45°C and about 65°C, although reverse phase high pressure liquid chromatography for most other polypeptides is typically carried out at room temperature or colder temperatures. Other useful restriction proteases (endoproteases) include thrombin, renin, and enterokinase, provided their recognition site has been engineered into the intervening spacer region of the fusion protein. Cyanogen bromide (CNBr) can also be used if a methionine intervenes between the peptide of interest and the rubredoxin

component, provided the peptide of interest contains no internal methionines that would result in undesired cleavage of the peptide upon contact with CNBr.

The invention further provides a method for making antibodies to a polypeptide of interest (i.e., a polypeptide antigen or hapten) using a rubredoxin fusion protein. A rubredoxin fusion protein comprising a rubredoxin polypeptide and the polypeptide antigen or hapten is introduced into a host, eliciting an immune response to the peptide antigen in a host cell. A cleavage site between the rubredoxin component and the fused polypeptide is not required as the rubredoxin moiety is negligibly antigenic. Thus, the fusion protein used in this method preferably does not contain a cleavage site. The method for making antibodies is not limited by the selection of a particular host; rather any desired host can be used such as a rabbit, goat, mouse, rat, cow or chicken. Antibodies are isolated and purified from the host using methods well-known in the art. The antibody is preferably a polyclonal antibody; however, the rubredoxin fusion protein can also be used to generate monoclonal antibodies to the polypeptide of interest.

The invention also provides a polypeptide vaccine comprising a rubredoxin fusion protein of the invention and a polynucleotide vaccine comprising a polynucleotide comprising a nucleotide sequence encoding a rubredoxin fusion protein. A preferred rubredoxin fusion protein for use in this embodiment of the invention includes, for example, a rubredoxin constituent linked to a polypeptide antigen or hapten. Preferably, the rubredoxin fusion protein used in or encoded by the vaccine is one wherein the N-terminal rubredoxin constituent is directly linked to the C-terminal fused polypeptide.

A vaccine is capable of generating an immune response in the animal to which it is administered. An immune response includes either or both of a cellular immune response or production of antibodies, and can include activation of the subject's B cells, T cells, helper T cells or other cells of the subject's immune system. Immunogenicity of rubredoxin fusion protein can be determined, for example, by administering the adjuvanted fusion protein to the subject, then observing of the associated immune response by analyzing antibody titers in the subject's serum.

In a preferred embodiment of the vaccine, the rubredoxin fusion protein used in the vaccine or encoded by the polynucleotide used in the vaccine further includes at least one epitope or epitope mimic, such as a T cell, helper T cell or B cell epitope or epitope mimic. Epitopes or epitope mimics can be derived from the species to which the vaccine is to be administered, from the species that was the source of the polypeptide antigen or hapten, or from any other species, including a virus, bacterium, or parasite. The use of immune cell epitopes derived from an immunogenic organism, such as a pathogenic parasite, is preferred.

A polynucleotide encoding a rubredoxin fusion protein can include DNA, RNA, a modified nucleic acid, or any combination thereof. The polynucleotide can be supplied as part of a vector or as a "naked" polynucleotide. General methods for construction, production and administration of polynucleotide vaccines are known in the art, e.g. F. Vogel et al., *Clin. Microbiol. Rev.* 8:406-410 (1995). Polynucleotides can be generated by means standard in the art, such as by recombinant techniques, or by enzymatic or chemical synthesis.

A polynucleotide used in a vaccine of the invention is preferably one that functionally encodes a rubredoxin fusion protein. A protein is "functionally encoded" if it is capable of being expressed from the genetic construct that contains it. For example, the polynucleotide can include one or more expression control sequences, such as *cis*-acting transcription/translation regulatory sequences, including one or more of the following: a promoter, response element, an initiator sequence, an enhancer, a ribosome binding site, an RNA splice site, an intron element, a polyadenylation site, and a transcriptional terminator sequence, which are operably linked to the coding sequence and are, either alone or in combination, capable of directing expression in the target animal. An expression control sequence is "operably linked" to a coding sequence if it is positioned on the construct such that it does, or can be used to, control or regulate transcription or translation of that coding sequence. Preferred expression control sequences include strong and/or inducible *cis*-acting transcription/translation regulatory sequences such as those derived from

metallothionine genes, actin genes, myosin genes, immunoglobulin genes, cytomegalovirus (CMV), SV40, Rous sarcoma virus, adenovirus, bovine papilloma virus, and the like.

The coding and expression control sequences for the rubredoxin fusion protein are preferably constructed in a vector, such as a plasmid of bacterial origin, a cosmid, episome, or a viral vector, for administration to a target animal. A vector useful in the vaccine of the present invention can be circular or linear, single-stranded or double stranded. There are numerous plasmids known to those of ordinary skill in the art useful for the production of polynucleotide vaccine plasmids. A specific embodiment employs constructs using the plasmid pcDNA3.1 as the vector (InVitrogen Corporation, Carlsbad, CA). In addition, the vector construct can contain immunostimulatory sequences (ISS) that stimulate the animal's immune system. Other possible additions to the polynucleotide vaccine constructs include nucleotide sequences coding cytokines, such as granulocyte macrophage colony stimulating factor (GM-CSF) or interleukin-12 (IL-12). The cytokines can be used in various combinations to fine-tune the response of the animal's immune system, including both antibody and cytotoxic T lymphocyte responses, to bring out the specific level of response needed to affect the animal's reproductive system.

Alternatively, the vector can be a viral vector, including an adenovirus vector, and adenovirus associated vector, or a retroviral vector. Preferably the viral vector is a nonreplicating retroviral vector such as the Moloney murine leukemia virus (N2) backbone as described by Irwin et al. (*J. Virology* 68:5036-5044 (1994)).

The polypeptide or polynucleotide vaccine is administered in a manner and an amount effective to cause the desired immune response in the animal. For example, a polypeptide vaccine can be administered in one or more doses, and typically includes between about 10 μg to about 2 mg of rubredoxin fusion protein. Likewise, a polynucleotide vaccine containing polynucleotide in an amount of about 5 μg to about 500 μg can be administered in one or more doses. One of skill in the art can readily determine a suitable dosage for a

particular animal, depending on the nature, size and overall health of the animal, as well as the condition to be treated.

A polypeptide or polynucleotide vaccine of the invention can be administered in any convenient manner. Forms of administration include intramuscular administration, subcutaneous or intradermal administration, oral administration, as by food or water, topical administration, including transdermal administration, aerosol administration, cloacal or vaginal administration, intracoelomic administration, intranasal administration, and transconjunctival administration, including the use of eye drops. In addition, liposome-mediated, microsphere-mediated, and microencapsulation systems are all included as delivery vehicles for the vaccine of the present invention.

Optionally the vaccine includes an adjuvant, the selection of which is a matter well-known to those of skill in the art and is influenced by the nature of the intended recipient.

EXAMPLES

The present invention is illustrated by the following examples. It is to be understood that the particular examples, materials, amounts, and procedures are to be interpreted broadly in accordance with the scope and spirit of the invention as set forth herein.

Example I.
Synthesis of a Rubredoxin Fusion Protein

Recombinant rubredoxin

Rubredoxins from numerous different organisms have been isolated, and the amino acid sequences of various rubredoxins and the genes encoding various rubredoxins have been published. In this experiment the gene encoding rubredoxin from *D. vulgaris* St. Hildenborough was used (see Fig. 1; also Bruschi et al., *Adv. Exp. Med. Biol.* 74:57-67 (1976); Voordouw, *Gene* 69: 75-83 (1988)). The gene was amplified by polymerase chain reaction (PCR)

from genomic DNA isolated from *D. vulgaris* using two primers and cloned into the expression vector pET24a (Novagen, Wisconsin) at the *Nde* I and *Bam*HI site. The pET-24a expression system utilizes the bacteriophage T7 promoter that serves as a binding site for T7 RNA polymerase and was incorporated into the chromosomal DNA of *E. coli* strain BL21 (DE3) (Novagen). T7 RNA polymerase is synthesized only upon the addition of isopropyl β-D-thiogalactoside (IPTG) to growing cultures since the gene for the T7 polymerase has been spliced into the chromosomal DNA of the *E. coli* host. The pET-24a plasmid also contains the gene for kanamycin resistance for selection of plasmid-containing colonies.

In the initial experiment, conditions were optimized for synthesis of rubredoxin in *E. coli*. Host cells were transformed and plasmid-containing colonies were obtained by kanamycin selection on Luria broth (LB), kanamycin plates. A single colony was transferred to 5 mL LB containing 50 ug/ml kanamycin (Sigma, St. Louis, MO) which was grown overnight at 37°C. The culture was then transferred to one liter of LB containing 50 ug/ml kanamycin and 100 uM $FeSO_4$ and grown to an optical density (OD_{590}) of 0.8 at 37°C. Induction of recombinant protein synthesis was initiated by the addition of 1 mM IPTG, after which the cells were allowed to grow for another 7-8 hours. Optimal incorporation of iron into the recombinant protein was obtained when the cultures were shifted to temperatures between 20-25°C after induction.

Construction of a recombinant rubredoxin fusion protein

To analyze whether a properly folded protein could be obtained if rubredoxin was fused at its C-terminus with another peptide region, a nucleotide sequence encoding the flag peptide (YKDDDDK, SEQ ID NO:13) affinity tag (IBI/KODAK, Rochester, NY), a polyhistidine sequence, and an enterokinase protease site was attached in frame at the C-terminal end of the rubredoxin gene, yielding pRUBEX1. The encoded peptide sequence provides two independent sites for affinity purification of the fusion protein along with a protease site for removal of the protein of interest from the fusion. Specifically, a resulting fusion protein can be purified by Ni-chelating affinity chromatography due to the

presence of the polyhistidine tag, and the flag peptide offers a second method for affinity purification using the monoclonal antibody-linked resin (IGM2) available from IBI Kodak.

All plasmids containing fusion constructs were transformed into *E. coli* strain BL-21, and the host cells were grown induced as described above for the rubredoxin optimization. After induction, the temperature was brought to 20°C for the final 7 hour growth period. Cells were harvested and stored at -70°C until needed. For expression of ^{15}N-labeled proteins and peptides, cultures were grown in M9 minimal media. Cells were initially streaked on M9 minimal media plates containing 50ug/ml kanamycin. A well-isolated colony was transferred to 100ml of M9 minimal media containing 1g/L ammonium-^{15}N chloride. The culture was grown at 37°C overnight. The 100ml innoculum (OD_{590}=3.0) culture was transferred to 900ml of M9 minimal media containing ammonium-^{15}N chloride as the nitrogen source (1g/L) supplemented with freshly prepared $FeSO_4$ for a final concentration of 30uM. At an OD_{590} of 0.7, additional $FeSO_4$ was added to bring the final concentration to 80uM. The cultures were induced with 1mM IPTG at an OD_{590} of 1 and were then transferred to 20°C and allowed to grow for an additional 15 hours. Cells were harvested and stored at -70°C until needed.

Cell disruption and Ni-chelating affinity chromatography

Frozen cell paste (12-15 grams, representing cells from 3 liters of media) was suspended in 100ml phosphate buffer (20mM, pH 7.4; 0.5M NaCl; Buffer A) and the resuspended cells were sonicated using a Branson Ultrasonic disrupter for 15 minutes (10 second pulses). The cell sonicate was spun at 10,000xg for 15 minutes and the supernatant which contained the soluble fusion protein was collected and processed as the cell-free extract.

High flow metal-chelating columns (5ml; Pharmacia) were used for purification of the fusion proteins. The column was washed and charged with 0.1M $NiSO_4$, washed again and then equilibrated with Buffer A containing 25mM imidazole. Imidazole was added to the cell-free extract to give a final concentration of 25 mM. This material was loaded onto the column at 3ml/min

and was washed with the equilibration buffer until the flow through was clear (4-6 bed volumes). The column was subsequently washed with 4 bed volumes of Buffer A containing 75mM and 150mM imidazole in order to elute several incomplete fusion products which were most likely formed as a result of incomplete translation. The complete fusion protein was finally eluted with Buffer A containing 300mM imidazole. Elution of the fusion proteins was monitored during purification by visual inspection of the column and flow through since the fusion products are red in color (due to the iron-sulfur center of rubredoxin). The purified protein (approximate volume 50-60 ml) was dialyzed overnight in 4 liter batches against a total of 12 liters of Tris HCl buffer (20mM, pH7.5). Total protein obtained after purification was estimated using the BCA assay (Pierce Biochemicals) with BSA as the standard.

The dialyzed fusion protein was brought to a concentration of 5-6 mgs/ml using an Amicon Centriprep (10K cut off) and was filtered using a 0.22μm syringe filter (Whatman) prior to storage in sterile falcon tubes. The protein keeps well at a concentration of 5-6mgs/ml at 4°C at pH 8.0 in the dark. Prolonged exposure to light (as in cold cabinets) leads to photobleaching of the protein and formation of a precipitate.

Analysis of the resulting fusion protein showed that fusion of the test peptide to the C-terminal end of rubredoxin did not alter any of the characteristics of the rubredoxin in terms of folding, ability to incorporate the non-heme iron center, protein yield, or protein thermostability. The final step involved the addition of a polylinker containing various restriction sites for the insertion of the gene sequences of the proteins, yielding pRUBEX3 (Fig. 1).

Example II.

Synthesis of Recombinant β-Amyloid Peptides as Fusions to Rubredoxin

Introduction

The β-amyloid 1-40 and 1-42 peptides (DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV[IA], SEQ ID NOS:10 and 14) generated by proteolytic cleavage of a membrane bound pre-protein (APP) represents a major constituent of the senile plaques which are deposited in the brains and cerebrovasculature of patients affected by Alzheimer's disease. The plaques are formed by ordered, self-aggregation of the peptides to form amyloid fibers. Onset of this disease is marked by enhanced levels of the longer and more hydrophobic $A\beta_{1-42}$ peptide in the brain (Iwatsubo et al., *Neuron* 13:45-53 (1994)); Lemere et al., *Nat. Med.* 2:1146-1150 (1996)); therefore, much attention is being directed towards determination of the tertiary structure of the monomeric peptides and higher order aggregates in an effort to find potential mechanisms of aggregation (Tomiyama et al., *Biochem. Biophys. Res. Commun.* 204: 76-83 (1994); Wood et al., *J. Biol. Chem.* 271:4086-4092 (1996)) and identify inhibitors of the process.

Any investigation that requires large quantities of peptide or protein necessitates the availability of a system that can be utilized to produce consistently pure working material. Currently, chemical synthesis of $A\beta_{1-42}$ is the main source of experimental material. Batch to batch variation in both quantity and polymeric state, the presence of truncated and blocked forms of the peptide, difficulty in separating incorrect synthesis products from the $A\beta_{1-42}$ peptide, and differing solubilities cause experimental results to differ among groups reporting aggregation results. Therefore, a method which would insure the production of pure, monomeric $A\beta_{1-40}$ and $A\beta_{1-42}$ peptides would greatly improve the consistency of results and would allow the use of methods which require large quantities of concentrated peptide such as Nuclear Magnetic Resonance (NMR) for structure determination. Labeling of peptides with the non-radioactive isotopes ^{15}N and ^{13}C greatly simplifies structural determination via NMR and would greatly benefit determination of structural changes that occur during the

aggregation process, but chemical synthesis of such labeled peptide is prohibitively expensive. Labeled peptides and proteins are easy to produce using recombinant techniques and are much less costly than those produced synthetically making this method very attractive to groups pursuing structural data.

Previous attempts to synthesize recombinant amyloid peptide in *E. coli* have resulted in the formation of inclusion bodies that required the use of guanidine thiocyanate for solubilization (B. Boyes et al., *J. Chromatog.*, 691:337 (1995); Gardella et al., *Biochem. J.* 294:667-674 (1993)). A method for synthesizing this peptide as a recombinant fusion protein occurring in inclusion bodies was previously developed at Hoffman-La Roche (Döbeli, et al., *Biotechnology* 13:988-993 (1995)), but processing of their fusion to form pure monomeric $A\beta_{1-42}$ is tedious in that it involves binding the fusion protein to a reverse-phase column followed by cyanogen bromide (CNBr) cleavage to remove the peptide from the fusion. Analysis of peptide purified with this method revealed formylation and carbamylation of the peptide as well as oxidation of Met-35. These alterations presumably occur as a result of CNBr cleavage of the peptide; Met-35 must be reduced by dimethylsulfoxide (DMSO) treatment in concentrated hydrochloric acid (HCl) before use. In this example, amyloid peptides were synthesized as fusions with rubredoxin in the hope of circumventing the difficulties of synthesizing homogeneous and consistently pure, monomeric peptides using existing methods. Recombinant synthesis as fusion proteins also allows more economical production of labeled peptides for use in continuing medical research efforts.

Accordingly, β-amyloid peptides 1-40 and 1-42 were synthesized as soluble recombinant fusion proteins using rubredoxin as a fusion partner. The fusion protein was purified by Ni-chelating chromatography and average yields of purified fusion product varied from 40-50 mg/L of culture. The fusion product was cleaved by restriction protease Factor Xa to separate the β-amyloid peptide from the rubredoxin carrier. The peptide was further purified by reverse phase chromatography at pH 6-8.5 at temperatures between about 45-65°C. The quality of the peptide was consistent from batch to batch and showed no

chemical modification as judged by mass spectrometric analysis. The purified peptides were biologically active and formed fibers at pH 2.5 as well as pH 6.5.

Construction of the expression vector

The DNA sequence encoding the β-amyloid 1-42 peptide was amplified by PCR using the human Alzheimers precursor protein (human βAPP) gene as template (provided by Dr. Sangram Sisodia, Johns Hopkins University, Boston, MA). During the PCR process (Bej et al., *Crit. Rev. Biochem. Mol. Biol.* 26:301-334 (1991)), a restriction protease site for Factor Xa was introduced at the amino terminal end of the β-amyloid 1-42 peptide for proteolytic cleavage from rubredoxin, in that the N-terminus primer designed and used for amplifying the β-amyloid 1-42 sequence encoded the residues Ile-Glu-Gly-Arg, the tetrapeptide recognition site for Factor Xa, along with a PstI restriction site (35 bases total). The C-terminus primer contained the sequence for the C-terminal region of the relevant peptide followed by a KpnI restriction site. The amplified DNA product was digested by PstI and KpnI and was ligated into the PstI-KpnI site of the polylinker region of pRUBEX3 (Example I) and sequenced. The final construct encoded a 13.6 kD fusion protein containing the rubredoxin gene, the His-Flag affinity site, the Factor Xa restriction site and the β-amyloid 1-40 or 1-42 peptides (Fig. 2). All constructs were initially made in pUC18, sequenced and then transferred into the expression vector pET24a at the Nde-BamH1 site.

Production of the rubredoxin-β-amyloid fusion protein

Expression of the fusion protein in one liter cultures was carried out essentially as described above in Example I. Expression in 20L fermentors were started by inoculating 50mls of overnight culture into a 24L fermentors containing 20L of LB supplemented with 100uM $FeSO_4$. The culture was grown at 37°C with stirring at 240 rpm and 3L of air/min. At an OD of 1.2 IPTG was added to a final concentration of 1mM and the temperature lowered to 20°C. Cultures were allowed to grow for another 6 hours. Cells were harvested and stored at -70°C. Average cell yield from a 20L fermentor run was 5.5-6g/liter.

Cells were disrupted, and Ni-chelating chromatography was carried out, substantially as described in Example I.

Digestion and cleavage of the fusion protein

Fusion amyloid protein was digested with Factor Xa (Boehringer-Mannheim) at a ratio (w/w) of 250:1 (fusion: protease) at room temperature overnight with continuous stirring. This procedure facilitated the aggregation of the cleaved amyloid peptides. The digest was finally centrifuged at 30,000xg for 30 minutes at 4°C. The aggregated peptide was collected as a pellet and was washed with water, 5mM EDTA to inhibit remaining Factor Xa, and finally by water before being stored at -20°C. This protocol enabled us to remove approximately 95% of the rubredoxin fusion partner and other soluble minor contaminants that might have co-purified with the fusion protein.

Purification of β-amyloid 1-42 and *1-40*

Following cleavage, the property of the $A\beta_{1-42}$ and $A\beta_{1-40}$ peptides to form sedimentable aggregates was used to concentrate and purify the peptide away from most of the rubredoxin moiety. But non-specific cleavage of both amyloid fusion proteins that occurs after Arginine-5 generated an additional peptide fragment that had to be separated from the intact peptides. The propensity of β-amyloid 1-42 to form aggregates and insoluble fibers poses a major problem in purifying this peptide (D. Burdick et al., *J. Biol. Chem.* 267:546 (1992), P. Sweeney et al., *Anal. Biochem.* 212:179 (1993)). Normal reverse phase chromatography is not a suitable method for purification. High temperature reverse phase chromatography using a Zorbax Stable Bond C18 column (McMod, PA) (B. Boyes et al., *J. Chromatog.* 691:337 (1995)) at pH 2.5 (0.05%TFA) was thus attempted. Temperatures in the range of 80-85°C resulted in good resolution between β-amyloid 1-42 and the various contaminating peaks. The β-amyloid 1-42 peptide isolated by this protocol was found to be pure as judged by mass spectrometry and was free of chemical modifications. However, this method poses a problem in that the temperatures used are very close to the boiling point of acetonitrile and further, heating a scale up preparatory column is

a long and expensive proposition. Moreover, it is difficult to work at a pH above about pH 6 with silica based resins since at high temperatures silica tends to degrade at a pH above 5.

It was discovered that separation was most readily achieved using a Vydac reverse-phase polymeric column with 5mM potassium acetate/5% acetonitrile (pH 8.0) as the aqueous phase and 5mM potassium acetate/10% isopropanol/80% acetonitrile (pH 8.0) as the mobile phase carried out at about 60°C. This polymer matrix produced good resolution at pH ranges of about 6 to about 8.5, and at temperatures between about 45 and about 65°C. Peptide recoveries were in the range of 65-80%. Separations were much sharper at 65°C than at 45°C, but the peak areas were very comparable at both temperatures indicating good recoveries. Low temperature purification is of further advantage since the stability and possible biohazards of subjecting peptides incorporating S^{35} methionine and the non-radioactive isotopes N^{15} or C^{13} to temperatures above 60°C are not known. Load capacity of a semi-preparative column in this material (10mm x 25cm) with good resolution of peaks was in the range of 100-200ugs of β-amyloid 1-42 peptide. It is expected that load levels in the range of 1.5-2mgs per run (25mm diameter X 25cm length) can be achieved. This would minimize loss in recovery of the peptide because of multiple runs and make the procedure much more economical.

Both peptides were judged to be completely intact and pure according to amino acid sequence results and mass spectrometry data after reverse phase separation on Vydac column. Mass spectrometric analysis of the molecular weight of several batches of peptide isolated from different fermentation runs by MALD-TOF and electrospray varied from 4514.6-4517.4 (expected MW = 4514.1) for the $A\beta_{1-42}$ peptide and 4328.2-4330.4 (expected MW = 4329.86) for the $A\beta_{1-40}$ peptide. The close agreement of the expected and actual weights clearly indicates that the peptides have not been chemically modified during any step of the purification protocol. The absence of additional peaks in the spectra indicates that the peptides are pure and reproducibility of the results from several fermentation runs shows that batch-to-batch peptide purity is maintained which is a major advantage over chemically synthesized peptide.

Mass spectrometric analysis of the $A\beta_{1-42}$ peptide purified from cells grown in minimal media containing ammonium-^{15}N chloride showed that the peptide was uniformly labeled with ^{15}N during expression, so production of labeled peptide is much more feasible and economical than chemical synthesis.

The most important biological assay for the amyloid peptides is their capacity to form fibers at room temperature. To circumvent the problem of the presence of pre-existing multimers which can form nuclei (act as seeds) for further aggregation in monomeric peptide solutions, we attempted fiber formation with $A\beta_{1-42}$ peptide freshly eluted (containing ~25% acetonitrile) from a reverse-phase column run at 65°C. The recombinant peptide was fully capable of forming fibers, as demonstrated by electron micrographs of fibers formed at pH 2.5 and pH 6.5 using peptide purified by this technique (not shown). Circular dichroism (CD) has also been used to show the consistent fiber-forming behavior of different batches of peptide.

Results

A soluble rubredoxin β-amyloid fusion protein was produced. The rubredoxin moiety folded correctly as judged by the successful incorporation of iron into the protein. The fusion protein was easily purified by Ni-chelating chromatography. Ni-chelating resins from several companies can be used (for example, Qiagen, Invitrogen and Boehringer Mannheim Biochemicals), but they do differ in binding and elution characteristics with respect to imidazole concentrations. The red color of the fusion provided a visible intrinsic marker to follow the protein during purification. Typical yields of the fusion protein were in the range of 40-50mgs/L as estimated by the BCA method. The fusion protein remained soluble at concentrations of 5-6mgs/ml at 4°C.

The average yield of β-amyloid 1-40 or β-amyloid 1-42 peptide was 3-4mgs/L. These recoveries can be further improved by employing larger columns and reducing the number of chromatographies to purify 3-4mg of peptide from 20 to one. Additionally, one of the main problems with expressing eukaryotic proteins in bacterial hosts is the altered bias in codon usage. By altering the codons of the eukaryotic gene to coincide with bacterial usage

(where feasible), it is probable that higher yields can be obtained. According to these data, decreasing the expression temperature may also lead to higher yields.

A major advantage of this recombinant system is the possibility of synthesizing radioactive peptides using S^{35}-labeled methionine. Purification of this peptide is possible at moderate temperatures of 45-50°C, conditions under which S^{35} is stable. Another advantage of this system is that it can be used for incorporating N^{15}, C^{13} and, with appropriate auxotrophs, various labeled amino acids into the β-amyloid peptides.

Example III.

Synthesis of the Extracellular Domain of Luteinizing Hormone Receptor (LHR) as a Fusion to Rubredoxin

Mass production of the extracellular domain of luteinizing hormone receptor (LHR) is of great commercial interest due to its potential for use as a contraceptive. Provided in the form of a "morning after" pill or other dosage form, extracellular LHR could act to prevent fertilization of the egg and/or uterine implantation of the fertilized egg.

Construction of the expression vector

The coding region of the *D. vulgaris* rubredoxin gene was cloned into the expression vector pET16b (Novagen), which contains a Factor Xa site at the appropriate location, at the *Xba*I/*Nco*I site to yield pRUBEX2. The amino-terminal 298 amino acid residues of human luteinizing hormone receptor (LHR), representing the extracellular domain, was then cloned into the *Nde*I/*Bam*HI site of pRUBEX2 to yield pRUBEX2-LHR. In pRUBEX2-LHR, a Factor Xa recognition site directly precedes the LHR coding region, and a spacer region is located between the rubredoxin coding region and the LHR coding region, thus including the Factor Xa site (Fig. 3). The spacer region further contains a poly-histidine tag to facilitate purification of the fusion protein. The total length of the spacer region (50 amino acids), which is longer than just the affinity sequence and the Factor Xa recognition sequence, was chosen to maximize the

efficiency of Factor Xa cutting to insure efficient separation of rubredoxin and LHR fragments after isolation of the fusion construct.

Expression of the rubredoxin-LHR fusion protein

Expression in one liter cultures were initiated by inoculating a single colony from a freshly streaked plate of pRUBEX2-LHR-transformed cells into 5ml of LB containing 50ug/ml kanamycin and growing the cells at 37°C for 8 hours. The culture was transferred into 1L of LB containing 100uM $FeSO_4$ and 50ug/ml of kanamycin and grown at 37°C in a gyratory shaker. The cultures were induced with 1mM IPTG when they reached an O.D. of 0.8 at 540nm and the temperature was lowered to 22°C. Cultures were grown for 8 hours, harvested by centrifugation and stored at -70°C until further use.

Purification of the rubredoxin-LHR fusion protein

Frozen cells (12-15g) were suspended in 100ml of 20mM phosphate buffer, pH 7.4, containing 0.5M NaCl (Buffer A) and sonicated in a Branson Ultrasonic Disrupter at full power for 15 minutes in 10 second pulses. The sonicate was centrifuged at 10,000 x g for 15 minutes and the supernatant which contained the fusion protein was used for further purification.

About 50ml of metal-chelating Sepharose (Pharmacia) was charged with 0.1M $NiSO_4$ and equilibrated with Buffer A containing 25mM imidazole. The column was washed with four bed volumes of equilibration buffer and then four bed volumes of equilibration buffer containing 150mM imidazole. The fusion protein was eluted with Buffer A containing 300mM imidazole. This process could be monitored by the intrinsic red color of the fusion protein. The purified protein was dialyzed overnight against three 4L changes of 20mM Tris-HCl pH 7.5. The fusion protein was then concentrated and washed with the Tris-HCl buffer using an Amicon Centriprep (10K exclusion) filter to remove all traces of imidazole, since imidazole is an inhibitor of Factor Xa protease.

Digestion and cleavage of the rubredoxin-LHR fusion protein.

Fusion protein was digested with protease Factor Xa (Boehringer Mannheim) at a ratio of 250:1 (fusion:protease) at 37°C for 45 minutes with constant stirring. This protocol resulted in the cleavage of about 95% of the fusion protein. The Xa-digested material was adjusted to 25mM imidazole and passed once again over the metal-chelating Sepharose resin. In this instance, the rubredoxin, which retained the poly-Histidine on its carboxy-terminus, bound to the resin while the LHR fragment passed through the column. The flow-through was successively dialyzed as follows: 1) for 3 hours in 1L of 50mM Tris-HCl pH 7.5, 10% glycerol and 1mM Cysteine; 2) for 3 hours in 1L of 50mM Tris-HCl pH 7.5, 10% glycerol, 1mM cysteine and 1mM cystine; and 3) overnight in 2L of 50mM Tris-HCl pH 8.0, 5mM DTT and 10% glycerol. The dialyzed material was concentrated and used for further experiments.

Results

The rubredoxin protein expression system produced 20-40mg/L of rubredoxin-LHR fusion protein. The fusion protein could be purified to greater than 95% purity by passage over a single Ni-Sepharose column. Although a second passage produced greater purity, it did not result in a more homogeneous LHR preparation and gave lower yields as would be expected.

The fusion protein was readily cleaved by low concentrations of Factor Xa provided that the Ni-Sepharose eluate had been thoroughly dialyzed to remove all traces of imidazole. Repassage over the Ni-Sepharose column resulted in the binding of all of the rubredoxin (and the red coloration); the LHR moiety was, in contrast, included in the flow through from the column. This step removed over 95% of the rubredoxin fusion partner and this purity could be improved by a second passage over the column with relatively small losses.

After dialysis, the recombinant LHR fragment cross reacted with LHR antibodies and could be used as an antigen for the production of polyclonal antibodies.

Although the rubredoxin moiety of the fusion folded properly as indicated by the binding of iron during folding and the red color of the protein,

the LHR moiety does not fold correctly as indicated by the failure to bind efficiently to luteinizing hormone (LH). However, it should be understood that the pRUBEX2 vector was not designed to produce a recombinant fusion protein that is secreted, and thus does not effect proper folding of some mammalian polypeptides that contain disulfide linkages. The extracellular domain of LHR, for example, contains at least four disulfide bonds; this apparently prevented it from folding to the native conformation in the reducing environment of the *E. coli* cytosol, a result which was not unexpected. On the other hand, rubredoxin-LHR fusion that is targeted for secretion would be expected to fold properly in the more oxidized periplasmic environment where the dsb protein, which is involved in disulfide bond formation and shuffling in *E. coli*, is present.

Example IV.

Rubredoxin Fusion Protein for the Generation of Polyclonal Antibodies

Construction of the expression vector

The coding region of the *D. vulgaris* rubredoxin gene was cloned into the expression vector pET21b (Novagen) at the *Nde*I/*Bam*HI site to yield pRUBEX1. A cDNA encoding the amino-terminal 340 amino acids of human luteinizing hormone receptor (hLHR), representing the extracellular domain (see Example III) was then cloned into the *Bam*HI site of pRUBEX1 to yield pRUBEX1-LHR which thus encodes a fusion protein consisting of the N-terminal extracellular domain of human LHR fused to the carrier protein rubredoxin at the C-terminal end of rubredoxin (Fig. 4).

Expression of the rubredoxin-LHR fusion protein

E. coli strain BL21 cells were transformed with pRUBEX1-LHR and a 1.0ml overnight culture of the transformed cells was inoculated into a 100ml culture and grown for 3 hours at room temperature prior to induction with 1mM IPTG for 3 hours at room temperature. Cells were collected at 5000 x g for 10 minutes and stored overnight at -20°C and then resuspended in 10ml of 50mMTris-HCl, pH 7.5 and disrupted with 5 second bursts of a sonicator at full

power until all cells were broken. The lysate was centrifuged at 25,000 x g for 15 minutes and the supernatant was discarded. The pellet was washed successively in water, 50mM Tris-HCl ph 7.5 containing 5mM EDTA, 50mM Tris-HCl pH 7.5, containing 5mM EDTA and 0.4% Triton X-100, water, and 50mM Tris-HCl, pH 7.5, containing 1mM EDTA. The washed pellet was solubilized in 5.0ml of 8M urea, and EDTA and phenylmethylsulfonylfluoride (PMSF) were added to final concentrations of 5mM and 2mM respectively. The solubilized protein was cleared for 30 minutes at 25,000 x g and then successively dialyzed as follows: 1) for 3 hours in 200ml of 50mM Tris-HCl, pH 7.5, 10% glycerol, and 1mM cysteine (Sigma Chemical Co., St. Louis, MO); 2) for 3 hours in 1L of 50mM Tris-HCl, pH 7.5, 10% glycerol, 1mM cysteine, and 1mM cystine (Sigma Chemical Co., St. Louis, MO); and 3) overnight in 1L of 50mM Tris-HCl pH 8.0, 5mM DTT, and 10% glycerol. The dialyzed material was fractionated in SDS polyacrylamide gels and a 3 kD band which was specific to transformed cells and immunoreactive with human LHR antibodies, was cut from preparative gels and washed with water. The excised bands were lyophilized, ground into powder and injected into rabbits. The initial injection was in Freund's complete adjuvant (Pierce Biochemicals) and was followed by three boosts in Freund's incomplete adjuvant (Pierce Biochemicals). Animals were bled and an IgG fraction was prepared from the serum.

Immunodetection of COS7-expressed rat LHR

Wild type rat LHR cDNA was cloned into pET24 to form pCDNA3. pCDNA3 and the empty vector pET24, as a control, were transiently transfected with lipofectamine into monkey kidney (COS7) cells grown in DMEM (Dulbecco's modified Eagle's medium; Gibco/BRL) supplemented with 10% fetal bovine serum (Gibco/BRL). Fifty hours following transfection, the cells were chilled on ice, washed with phosphate buffered saline, and extracted in 150mM NaCl, 20mM HEPES pH 7.4 (Sigma Chemical Co., St. Louis, MO) and 0.5% Nonidet-P40 (Sigma Chemical Co., St. Louis, MO) in the presence of 0.5mM N-ethyl maleimide Sigma Chemical Co., St. Louis, MO), 0.2mM PMSF and 0.5mM EDTA. Cells were incubated in the extraction buffer for 20 minutes

on ice and the solubilized fraction was separated by centrifugation at 13000 x g for 10 minutes. The native or denatured cell extracts (~20ug protein) were incubated with N-glycosidase F (0.6 units) for 1 hour at 37°C. Cells extracts were denatured in 1% SDS for 5 minutes at 100°C and then diluted to 0.1% SDS for subsequent procedures. The products were reduced with β-mercaptoethanol (Sigma Chemical Co., St. Louis, MO) and fractionated on a 10% SDS-polyacrylamide gel and then transferred to nitrocellulose membrane. The blot was blocked with 2.5% bovine serum albumin (BSA) and incubated for 12 hours with a rabbit anti-hLHR antibody. Chemiluminescent immunodetection was performed employing the ECL system from Amersham Co. (Arlington Heights, IL).

Results

The induced rubredoxin-LHR fusion protein was readily visible by Coomassie Blue staining after fractionation of whole bacterial cell lysates from transformed BL21 *E. coli* cells in SDS polyacrylamide gels. Large amounts of the fusion protein were produced; estimates from the stained gels suggest that from 10-20mg of fusion protein was produced in 500ml of cells. The fusion protein was easily centrifuged from cell lysates, but it was also readily soluble in 8M urea.

Fusion protein bands excised from SDS-polyacrylamide gels and ground into a fine powder were excellent antigens in New Zealand which rabbits. Although three boosts were administered before bleeding the animals, it is not known if they were all necessary. The IgG fraction purified from the sera of inoculated rabbits did not react with native human LHR or rubredoxin, but reacted only with the recombinant hLHR fusion protein or deglycosylated native hLHR. As the fusion protein expressed in *E. coli* that was used for antigen is not glycosylated, it is not surprising that antisera directed against the fusion protein did not react with native hLHR which contains six known N-linked glycosylation sites. When these carbohydrates were stripped from the native protein, however, the antisera cross-reacted with the human protein. It was surprising, on the other hand, that the antisera did not cross-react with native or denatured rubredoxin, as

the rubredoxin comprised about 50% of the fusion protein. Attempts in our laboratory to make rabbit polyclonal antibodies to *D. vulgaris* rubredoxin have been unsuccessful, however, suggesting in combination with these results that rubredoxin may fortuitously be a very poor antigen. Even when rubredoxin was added to gels in extremely high concentration, we were unable to elicit a cross-reaction with the IgG fraction purified from the sera of rabbits innoculated with the fusion protein.

In order to detect the expression of rat LHR glycoproteins in transfected COS7 cells, the cells were lysed and deglycosylated with N-glycosidase F. Polyclonal antibodies elicited in rabbits with the recombinant rubredoxin-LHR fusion protein cross-reacted with proteins of 62 and 40kD after fractionation of the deglycosylated COS7 proteins. The smaller protein is most likely a degradation product of the 62kD protein generated by the unmasking of protease sites during the oligosaccharide modifications.

Example V.

Synthesis of a Pig Leptin/Rubredoxin Fusion Protein

Leptins are 12-15 kDa proteins which are known to be involved in the regulation of obesity in humans and other mammalian organisms. Expression of various leptins (human, rat, mouse and pig) by themselves or as fusions in *E. coli* have invariably led to the formation of inclusion bodies (K. Giese et al., *Mol. Med.* 2: 50-58 (1996); A. Fawzi et al., *Horm. Metab. Res.* 28: 694-697 (1996)). The inclusion bodies can be resolubilized and the proteins refolded to yield active leptin with varying degrees of success. Our own attempts to purify over-expressed pig leptin led to extremely poor recovery of active protein after the final re-folding step. A rubredoxin/pig-leptin fusion was therefore constructed to assess whether soluble leptin fusion protein could be produced that would yield a greater amount of recoverable, active leptin.

Construction of pig leptin/rubredoxin fusion

Native pig leptin contains a 21 amino acid signal peptide which is absent in the mature processed protein. In designing the rubredoxin fusion construct, this signal peptide sequence was deleted so that the amino-terminus originated at Val-22 of the pre-leptin sequence. The N-terminus primer was designed according to the amyloid protein scheme and included the KpnI restriction site and a Factor Xa recognition site just before the initial residues of the leptin sequence. The C-terminus primer contained the sequence for the C-terminal region of the protein along with a HindIII restriction site. The gene was synthesized by PCR amplification using a cDNA clone as the template. The amplified product was digested with KpnI and HindIII and was ligated into the corresponding site of pRUBEX 3 (Example I). After transformation of the plasmid into *E. coli* DH5α (strain BL-21 as described in Example I), three recombinant clones were isolated and determined by restriction analysis to contain the entire fusion protein gene.

Purification, digestion and cleavage of the leptin/rubredoxin fusion protein

The fusion protein was purified as described in Examples I and II, and the yield of the soluble leptin fusion was about 10-15mgs/liter. Leptin fusions were digested with Factor Xa at a ratio (w/w) of 100:1 at pH 8.0 at room temperature. Leptin fusions were also digestible with recombinant enterokinase, but not with native enterokinase. The digest was centrifuged at 15,000xg for 15 minutes and the supernatant was used for analysis.

Analysis of the purified leptin/rubredoxin fusion protein

The purified fusion protein and the Factor Xa digests were analyzed on a 10% Tris-tricine/sodium dodecyl sulfate (SDS) polyacrylamide gel (see Fig. 5). Lane 4 shows purified, undigested fusion protein (arrow; 22 kD, 5μg) but the mobility of the band is retarded due to the presence of the histidine moiety due to its positive charge; lanes 5 and 6 show a 7-hour digest of fusion protein (15μg and 10μg, respectively) with Factor Xa. The 14 kD band (top arrow) represents pure leptin and the 9.3 kD band (bottom arrow) represents the

rubredoxin-histidine portion of the fusion just before the Factor Xa site. Again, the mobility of the 9.3 kD band is retarded due to the presence of the histidine moiety. The presence of leptin in the supernatant indicates that leptin is soluble after digestion with the protease. These results were confirmed by western-blot analysis (Fig. 6). Lane 1 contains pig leptin fusion protein digested with Factor Xa (150ng); Lane 2 contains purified pig leptin fusion protein (100ng). The membrane was exposed to fluorescent-labeled antibody raised against purified pig leptin. The two lanes shown in Fig. 6 were cross-reacted to antibody raised against pig leptin. Both of the products cross-reacted with the antibody thereby indicating the presence of leptin in the fusion and in the digested fusion.

Example VI.

Synthesis of Feline Pro-Insulin/Rubredoxin Fusion Protein

Recombinant pro-insulin synthesized in *E. coli* is the major source of pharmaceutical grade insulin used in the treatment of diabetes. *In situ*, insulin is initially produced as a pro-insulin chain composed of three domains, A, B, and C, which contain two intramolecular disulfide bonds. During maturation of the protein, domain C is cleaved from the A and B domains and the result is a heterodimeric insulin molecule whose two subunits are joined by two disulfide bonds. *In vitro*, two strategies have been employed for the synthesis of mature insulin. One strategy involves reconstitution of the separately synthesized subunits, A and B, to form active insulin while the second strategy involves synthesizing the pro-insulin (all three domains) as an insoluble single chain in inclusion bodies. After successful solubilization and refolding of the pro-insulin, subunit C is removed by cleavage with trypsin and carboxypeptidase C to yield active insulin. The latter method has been reported to give significantly higher levels of active insulin, although pro-insulins from different animal sources have different intrinsic solubilities. A feline pro-insulin/rubredoxin fusion was therefore constructed in order to more efficiently recover soluble fusion protein.

Construction of the feline pro-insulin/rubredoxin fusion

The gene encoding feline pro-insulin was synthesized as constituent oligonucleotides which were ligated together to form a single composite gene. The codons were altered according to an *E. coli* codon usage table to maximize expression. A Factor Xa site was included at the 5' end of the pro-insulin oligonucleotide containing the N-terminus sequence. The composite gene was then digested with KpnI-HindIII and was ligated into the corresponding sites of RUBEX 3 and finally transformed into *E. coli* DH5α (strain BL-21 as described in Example I). Several recombinant clones were isolated and sequenced to verify the complete incorporation of all portions of the sequence.

Purification, digestion and cleavage of the feline pro-insulin/rubredoxin fusion protein

The fusion protein was purified as described in Examples I and II, and the yield of the soluble pro-insulin fusion was about 25 mgs/liter. Pro-insulin fusions were digested with Factor Xa at a ratio (w/w) of 100:1 at pH 8.0 at room temperature. Pro-insulin fusions were also digestible with recombinant enterokinase, but not with native enterokinase. Digestion with enterokinase reduced the amount of non-specific cleavage products compared to Factor Xa.

Analysis of the purified feline pro-insulin/rubredoxin fusion protein

The rubredoxin pro-insulin fusion migrated as a 19 kD band on a 10% Tris-Tricine native gel (Figure 7, lane 1, 5μg). Digests of the fusion with Factor Xa showed a number of non-specific cleavage products; therefore, digestion with recombinant enterokinase (an enterokinase site being a part of the flag peptide sequence) was attempted. The fusion protein was digested at a w/w ratio of 75:1 (fusion:enzyme) overnight at room temperature. The digest was centrifuged at 15,000 x g for 20 minutes and the supernatant was analyzed on a 10% Tris-Tricine gel. The digest revealed the two expected bands: a 9.6kDa rubredoxin band (top arrow) and a 9 kD pro-insulin band (bottom arrow; Figure 7, lanes 2 and 3, 10μg and 7μg, respectively). The mobility of the 9.6 kD band was retarded due to the presence of the histidine moiety. The 9 kD band (bottom

arrow) was electrophoretically transferred to a PVDF membrane and was analyzed via amino acid sequencing. The first twenty amino acids were determined and were found to match the expected sequence of pro-insulin, except for an additional portion of the polylinker which was present as a result of the location of the enterokinase restriction site in the fusion protein.

Sequence Listing Free Text

(SEQ ID NO:1) portion of pRUBEX

(SEQ ID NO:2) modified rubredoxin including affinity tag, flag peptide and enterokinase site

(SEQ ID NO:4) affinity tag

(SEQ ID NO:5) Flag peptide

(SEQ ID NO:6) enterokinase site

(SEQ ID NO:7) affinity tag

(SEQ ID NO:8) $A\beta_{1-42}$ rubredoxin fusion construct

(SEQ ID NO:9) $A\beta_{1-42}$ rubredoxin fusion protein

(SEQ ID NO:10) $A\beta_{1-42}$ peptide

(SEQ ID NO:11) Factor Xa restriction site

(SEQ ID NO:12) intervening spacer region

(SEQ ID NO:13) Flag peptide

(SEQ ID NO:14) $A\beta_{1-40}$ peptide

The complete disclosure of all patents, patent applications, database information (e.g., electronically available GenBank amino acid and nucleotide sequence submissions) and publications cited herein are incorporated by reference. The foregoing detailed description and examples have been given for clarity of understanding only. No unnecessary limitations are to be understood therefrom. The invention is not limited to the exact details shown and described, for variations obvious to one skilled in the art will be included within the invention defined by the claim.